

## 論文の内容の要旨

論文題目 進行非小細胞肺癌患者における QOL 評価構造の個人間差に関する研究  
—ランダム係数モデルを用いた検討—

氏 名 森田智視

### 1. はじめに

癌など治癒が望みにくい疾患においては、治療法の選択がその後の Quality of Life (QOL) に大きな影響を与えるため、そのような場面では治療法の有効性と安全性に関する客観的なエビデンスだけでなく QOL に関する情報も提供されることが重要である。現在、QOL の画一的な定義は存在しないものの、QOL を抽象的な構成概念として、活動性、身体性、精神性、社会性などの複数の次元 (ドメイン) から構成され、その評価は主観的なものであることについてはコンセンサスが得られている。さらに、QOL は次の三つの水準、1) 最上位の水準: 総合的な QOL (グローバル QOL)、2) 中間の水準: 複数のドメイン、3) 最下部の水準: 各ドメインの構成要素である QOL 調査票の質問項目、から構成されると説明されている。患者個々人が最上位にあるグローバル QOL を評価する際には、各ドメインが各個人にとってどれほど重要であるかという主観的な評価に基づいて多くの判断を行う必要がある。つまり、各個人のグローバル QOL におけるそれぞれのドメインの相対的な重要性を反映している“重み”に基づいて全ドメインを統合したものがグローバル QOL であると考えられる。従って、重みは患者間で異なることが予想され、また、治療中と治療後のように患者がおかれる状況の違いによって変化するかもしれない。

これまで、重みに関する研究では、グローバル QOL および複数の QOL ドメインをそれぞれ目的変数および説明変数とした古典的な線形回帰分析が用いられ、患者集団全体での重みについて検討されてきた。しかしながら、各ドメインに対する重みが患者個人間で異なるのかどうか、さらにその個人間でのバラツキが癌化学療法の実行中と実行後で異なるかどうかを検討した研究はこれまで行われていない。

本研究では、進行非小細胞肺癌患者を対象とし、ランダム係数モデルを用いて QOL の各ドメインに対する“患者個人ごとの重み”および“集団平均としての重み”を推定することにより QOL 評価の構造を調べた。すなわち、1) 各ドメインに対する個人ごとの重みは患者間で大きくばらつくか、2) 各ドメインに対する集団平均としての重みは有意に大きいのか、すなわち各ドメインはグローバル QOL に関連するか、3) それら重みの構造は化学療法実行中と実行後で異なるか、の三つを調べることを目的とした。なお、

QOL 評価の構造を検討するにあたり、本研究で用いた QOL 調査票である QOL Questionnaire for Cancer Patients Treated with Anticancer Drugs (QOL-ACD) の進行非小細胞肺癌における信頼性と妥当性を確認した。

## 2. 対象と方法

進行非小細胞肺癌患者（臨床病期 IIIB または IV、ECOG Performance Status (PS) 0~2）を対象として実施された癌化学療法の一つの第 III 相無作為化比較臨床試験に登録された患者のうち、QOL 調査に協力の得られた症例を研究対象とした。二つの試験は、共通のエンドポイント、投与方法および評価スケジュールを用いて三つの治療法を比較することを目的として実施された。全ての治療群において 1 コースを 4 週間として最大 6 コースまで治療が繰り返され、各コース第 1、8、15 日目に抗癌剤の投与が行われた。QOL-ACD は、全 22 項目の自記式 QOL 調査票であり、活動性（項目 1~6）、身体性（項目 7~11）、精神性（項目 12~16）、社会性（項目 17~21）の 4 ドメインと、グローバル QOL を評価するためのフェーススケール（項目 22）から構成されている。QOL の調査は、治療前、治療中（毎週各コース第 1、8、15、22 日の前）および治療後 1 ヶ月間（毎週）に行った。

QOL-ACD の信頼性と妥当性については、治療前に調査票記入のあった症例を解析対象とし、1) 計量心理学的解析と臨床的判断による不適当な項目の除外、2) QOL-ACD のドメインの確認と臨床的妥当性の検討、の二つのステップで評価を行った。QOL ドメインの重みについては、治療前および治療開始後に調査票記入のあった症例を解析対象とし、古典的な線形回帰分析を拡張し回帰係数の個人間差を考慮したモデルであるランダム係数モデル（次式）を用いて各ドメインに対する重みの個人間差を調べた。なお、ランダム係数モデルの目的変数および説明変数にはそれぞれフェーススケールと 4 ドメインのスコアの治療前からの変化量を用いた。

$$Y_i = X_i \beta + X_i b_i + e_i$$

$$\begin{pmatrix} b_i \\ e_i \end{pmatrix} \stackrel{i.i.d.}{\sim} N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} D & 0 \\ 0 & \Sigma_i \end{pmatrix} \right]$$

ただし、 $Y_i$  は患者  $i$ （患者  $i=1, \dots, n$ ）に対する目的変数ベクトル、 $X_i$  は切片およびドメインを表す説明変数行列である。 $\beta$  は切片およびドメインに対する集団平均の回帰係数を含むパラメータベクトル、 $b_i$  は患者  $i$  における切片および回帰係数を含むランダム係数ベクトルであり、 $e_i$  は患者  $i$  に対するランダム誤差ベクトルである。 $D$  および  $\Sigma_i$  はそれぞれ  $b_i$  および  $e_i$  の分散共分散行列を表す。この時、個人特有の回帰係数  $b_i$  は母集団でのパラメータの非一様性を反映しており、平均的な回帰係数からどれだけ各個人が逸脱しているかという逸脱の程度として解釈することができる。本研究では、患者ごとに計算される回帰係数を個人ごとの重みとし、それらの平均値を集団平均の重みとした。各ドメインに対する個人ごとの重みが患者間でばらつくかどうかを調べることは、 $D$  の構造を決定することに対応する。すなわち、各ドメインに対する個人ごとの重みの分散が 0 であるかどうかを調べればよい。しかしながら、負の値をとることのない分散成分の有意性検定については、帰無仮説（分散成分=0）が対立仮説のパラメータ空間の境界に存在するため、通常の尤度原理が適用できない。そこで、帰無仮説のもとで対応する尤度比検定統計量が、二つの  $\chi^2$  分布の混合分布に漸近的に従うことを利用した。また、治療中と治療後で重みの構造を比較するため、時期ごとに別々に解析を行った。さらに、QOL 評価脱落例の結果に与える影響を調べるため、治療中については、治療前、中、後の全てにおいて調査票記入を行った患者で構成されるサブセットの患者集団でも解析を行った。

### 3. 結果

1995年6月から1998年1月までに計583例の適格例が登録され、治療前の調査票記入に協力の得られた395例においてQOL-ACDの信頼性・妥当性を調べた。入浴に関する項目6はtest-retest再現性に問題があり、心の支えに関する項目16は収束妥当性に問題があった。項目6と項目16は計量心理学的および臨床的観点から不適切であると判断し、それらをドメインスコアの計算から除いたところ、全てのドメインにおいて高い内的整合性とtest-retest再現性が得られ、調査票開発時に想定されたドメイン構造が確認された。また、各ドメインとPSや体重減少などの臨床パラメータとの間に有意な関係が見られた。

QOLドメインの重みについては、治療前および治療中に調査票記入が行われた377例を解析対象とした。ランダム係数の分散成分に対する検定を行い有意となったドメインの結果を表1に示す。

表1. ランダム係数モデルに含まれた個人ごとの重み

| 治療中                 |      |                 | 治療後                        |      |                 |                     |      |                 |
|---------------------|------|-----------------|----------------------------|------|-----------------|---------------------|------|-----------------|
| 全解析対象 (n=377)       |      |                 | サブセット <sup>a</sup> (n=223) |      |                 | (n=223)             |      |                 |
| ドメイン                | 推定値  | p値 <sup>b</sup> | ドメイン                       | 推定値  | p値 <sup>b</sup> | ドメイン                | 推定値  | p値 <sup>b</sup> |
| 切片                  | 0.31 | <0.001          | 切片                         | 0.30 | <0.001          | 切片                  | 0.41 | <0.001          |
| 身体性                 | 0.06 | <0.001          | 身体性                        | 0.06 | <0.001          | 活動性                 | 0.05 | 0.013           |
| 社会性                 | 0.12 | 0.016           | 社会性                        | 0.15 | 0.013           | 社会性                 | 0.20 | <0.001          |
| ランダム誤差              | 0.20 | -               | ランダム誤差                     | 0.22 | -               | ランダム誤差              | 0.16 | -               |
| あてはまり度 <sup>c</sup> | 0.45 |                 | あてはまり度 <sup>c</sup>        | 0.47 |                 | あてはまり度 <sup>c</sup> | 0.50 |                 |

<sup>a</sup>治療前、中、後全てにおいて調査票記入を行った患者 <sup>b</sup>ランダム係数の分散成分に対する混合 $\chi^2$ 尤度比検定

<sup>c</sup>モデルのあてはまり度 =  $1 - \frac{\text{重みを含むモデルの切片ランダム係数とランダム誤差の分散の和}}{\text{重みを含まないモデルの切片ランダム係数とランダム誤差の分散の和}}$

表2. 集団平均の重み

| ドメイン | 治療中           |      |        |                            |     |        | 治療後     |     |        |
|------|---------------|------|--------|----------------------------|-----|--------|---------|-----|--------|
|      | 全解析対象 (n=377) |      |        | サブセット <sup>a</sup> (n=223) |     |        | (n=223) |     |        |
|      | 推定値           | t値   | p値     | 推定値                        | t値  | p値     | 推定値     | t値  | p値     |
| 活動性  | 0.12          | 4.3  | <0.001 | 0.10                       | 2.8 | 0.006  | 0.08    | 1.5 | 0.151  |
| 身体性  | 0.41          | 12.9 | <0.001 | 0.39                       | 9.5 | <0.001 | 0.27    | 4.2 | <0.001 |
| 精神性  | 0.18          | 5.0  | <0.001 | 0.20                       | 4.3 | <0.001 | 0.22    | 3.0 | 0.004  |
| 社会性  | 0.21          | 4.5  | <0.001 | 0.24                       | 4.0 | <0.001 | 0.34    | 4.1 | <0.001 |

<sup>a</sup>治療前、中、後全てにおいて調査票記入を行った患者

治療中、有意に大きなバラツキが見られたのは、身体性と社会性のドメインであった。同様の結果がサブセット(223例)における解析からも得られた。治療後、有意に大きなバラツキが見られたのは、活動性と社会性のドメインであった。次に、集団平均の重みを表2に示す。治療中は、全解析対象377例およびサブセットの223例において4ドメイン全ての重みが有意であり、身体性ドメインのt値が最大であった。治療後は活動性を除く3ドメインの重みが有意であり、t値の大きさは同様であった。

### 4. 考察

QOL-ACDの進行非小細胞肺癌における信頼性と妥当性を調べたところ、項目6は入浴スケジュールが管理されてしまう入院下で再現性に問題を生じ、項目16は四つの何れのドメインとも異なる側面を測定している可能性が高いと考え、2項目をドメインスコアの計算から除いた。その結果、QOL-ACDは計量心理学的にも臨床的にも妥当な調査票であることが確認された。2項目を除いた後、ランダム係数モデルを用いてグローバルQOLと4ドメインとの関係を解析することによって個人ごとの重みの個人間差を調べ

た。ランダム係数モデルは、これまで重みの評価に用いられてきた古典的線形回帰分析を拡張し、回帰係数の個人差を考慮したモデルである。本研究では、個人特有の回帰係数は個人間で互いに独立に同一の確率分布に従う確率変数であると仮定し、その分布には数学的にも都合がよく推定手順が容易になる正規分布を用いた。ランダム誤差についても同様に正規分布を仮定した。それらの仮定を確かめるため、個人特有の回帰係数とランダム誤差の推定値を2種類の残差とし、それらのヒストグラムを別々に調べた。その結果、何れの残差についても極端に大きく外れた値もなくほぼ左右対称な分布が得られたことにより、正規分布の仮定が概ね満たされたと判断した。また、個人ごとの重みと集団平均の重みによって目的変数であるフェーススケールの約50%の変動が説明されており、重みとしての解釈が十分可能であると考えられた。以上より、本研究であてはめたモデルは適切であると判断した。

治療中、身体性と社会性のドメインに対する重みに大きな個人間差がみられたのは、前者は化学療法による有害事象の受け止め方の違いに、後者は抗癌治療に関する経済的な問題の捉え方の違いや多様な患者-家族関係に起因しているのではないかと推察された。治療後の社会性ドメインに対する重みの個人間差は、治療結果や病状だけでなく家族のことを含めた将来の社会生活に対する不安の感じ方が多様であったためであり、活動性の重みについては、治療終了後の退院により日常生活の活動的側面に対して集団全体としては関心が薄れたが、その重要性は個人間で大きく異なったためではないかと推察された。重みに患者間差があることが分かればそれに応じて治療方針を決定することが重要となる。例えば、ある患者がスコアを最も低く（悪い状態）かつ最大の重み（最も重要）をつけたドメインに対して集中的にケアを行うことによって、その患者のグローバル QOL が著しく改善することが期待される。したがって、個人ごとの重みの評価は、日常診療において患者個人に対してどのようなケアを行うべきかを決定する際の重要な判断材料の一つになる可能性がある。集団平均の重みについては、治療中、身体性ドメインが最も強くグローバル QOL に関連した。身体性ドメインが集団全体として重要であったのは、治療中に発現した悪心・嘔吐や食欲不振などの有害事象により身体面が大きく影響を受けたためではないかと考えられた。このことは、化学療法施行中の QOL を維持するためには、抗癌剤副作用に対する症状コントロールを全ての患者に対して標準的に行うべきであることを示唆しているのかもしれない。

本研究でドメインに対する重みの解析対象となったのは、臨床試験の適格症例 583 例のうち QOL 調査に協力の得られた 377 例であった。解析対象患者と除外患者の背景因子にわずかながら違いが見られたため、得られた結果の一般化可能性には問題があるのかもしれない。さらに、治療後の解析からは、8 例の死亡を含む QOL 評価脱落あるいは積極的な後治療開始のため 154 例を除外した。したがって、治療後の解析結果にはバイアスが生じている可能性があり、また、治療中と治療後の重み構造の比較には注意を要する。しかしながら、治療中の重みの構造に関して、治療中および治療後ともに解析対象となったサブセット 223 例と全解析対象 337 例において同様の結果が得られたことは、サブセット内での治療中と治療後の比較は妥当であることから、本研究で得られた結果の妥当性を十分支持するものであると思われる。

## 5. 結論

QOL-ACD は、進行非小細胞肺癌患者の QOL 評価において計量心理学的に妥当であり、臨床的にも有用な調査票であることが確認された。ランダム係数モデルを用いることによって、治療中は身体性と社会性のドメインに対する個人ごとの重みが患者間で大きくばらつき、集団平均として四つのドメイン（特に身体性ドメイン）がグローバル QOL に関連することが示された。治療後は、活動性ドメインと社会性ドメインの個人ごとの重みが患者間で大きくばらつき、集団平均として活動性以外のドメインがグローバル QOL に関連することが分かった。本研究により日常診療における治療の個別化のために各ドメインに対する重みを QOL スコアとともに評価することの重要性が示唆された。